Bringing Order to Chaos: ScienceBase and Other Project Lifecycle Tools

September 5, 2013

Emily Fort
Data and Information Coordinator
USGS National Climate Change & Wildlife Science Center
Reston, VA

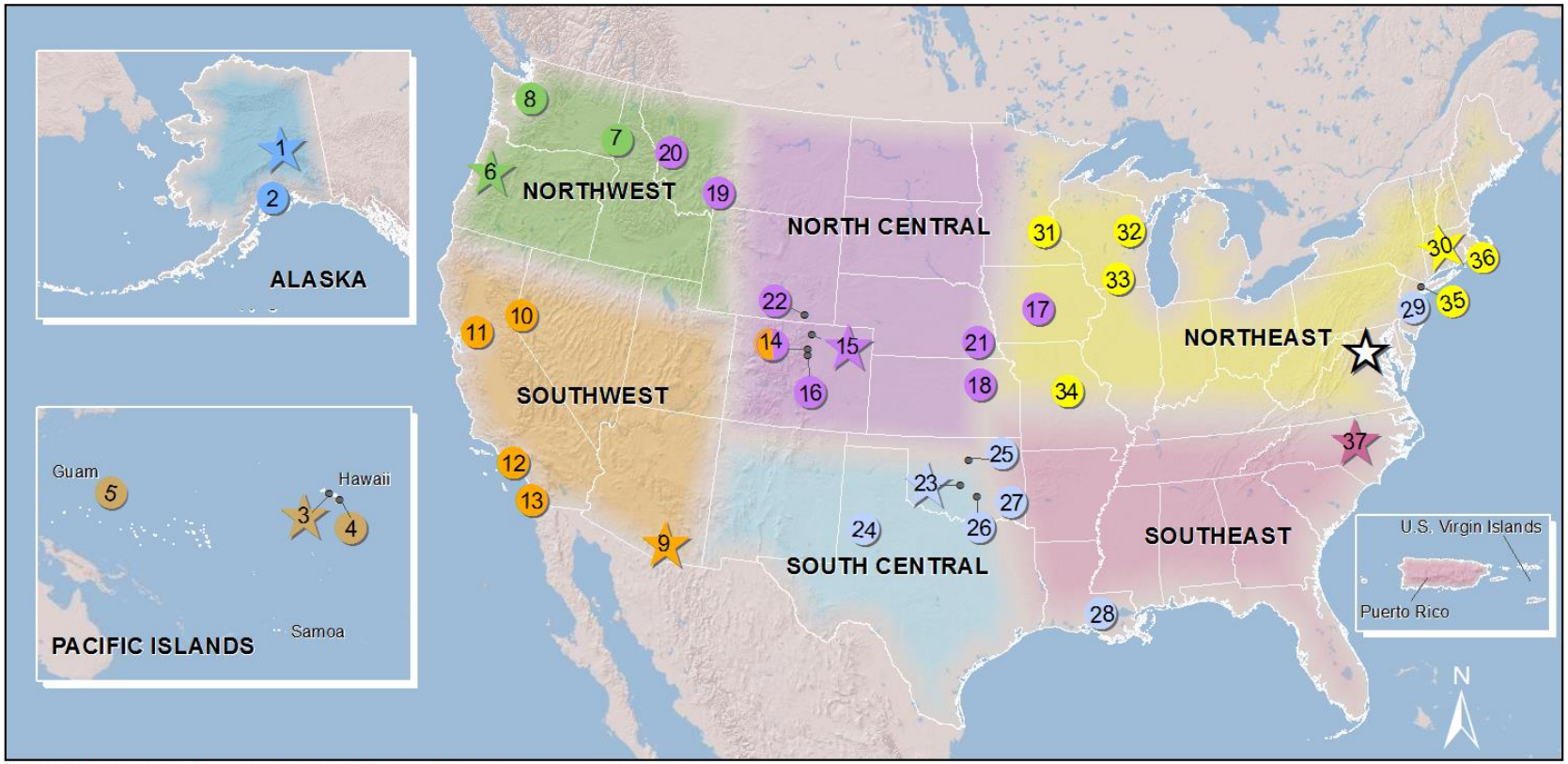
USGS
science for a changing world

The Plan

- Who we are, challenges, opportunities
- Approach, policies, process, big picture
- Pieces of the big picture
- Integration
- The future

How best to support NCCWSC, the CSCs, data management and provide some core tools?

USGS science for a changing world



Base from ESRI, 2009, Albers Equal Area Conic Projection, North American Datum of 1983

EXPLANATION

National Climate Change and Wildlife Science Center

CSC Lead Institutions

CSC Institutions

Alaska CSC
1. University of Alaska - Fairbanks
2. University of Alaska - Anchorage

Pacific Islands CSC
3. University of Hawaii at Manoa
4. University of Hawaii at Hilo
5. University of Guam

Northwest CSC
6. Oregon State University
7. University of Idaho
8. University of Washington

Southwest CSC
9. University of Arizona
10. Desert Research Institute (Nevada)
11. University of California - Davis
12. University of California - Los Angeles
13. Scripps Institute of Oceanography
14. University of Colorado

North Central CSC
14. University of Colorado
15. Colorado State University
16. Colorado School of Mines
17. Iowa State University
18. Kansas State University
19. Montana State University
20. University of Montana
21. University of Nebraska - Lincoln
22. University of Wyoming

South Central CSC
23. University of Oklahoma
24. Texas Tech University
25. Oklahoma State University
26. Chickasaw Nation
27. Choctaw Nation of Oklahoma
28. Louisiana State University
29. NOAA Geophysical Fluid Dynamics Laboratory

Northeast CSC
30. University of Massachusetts Amherst
31. University of Minnesota
32. College of Menominee Nation
33. University of Wisconsin - Madison
34. University of Missouri Columbia
35. Columbia University
36. Marine Biological Laboratory

Southeast CSC
37. North Carolina State University

Challenges

- Need to know what we are doing and where across the network
- Need to provide some basic capabilities and consistency for the NCCWSC and the CSCs
- Ensure data policies are enforced
- Provide access for university + federal scientists
- Fund ~100 projects each year
- Recognize reality that we are a new and growing program with a small staff

USGS
science for a changing world

Opportunities

- Blank slate
- Government policy
- Supportive management
- Recognition of importance of tools and data management
- Extensive capabilities at USGS, other federal agencies, and at the universities

Approach

- Don’t build another stove pipe
- Provide support to CSCs – both tools **and people**
- Develop data policies
- Link projects to data and products
- Use standards and web services
- Identify integration opportunities
- Develop core capabilities that our partners can link to, integrate, re-imagine

USGS
science for a changing world

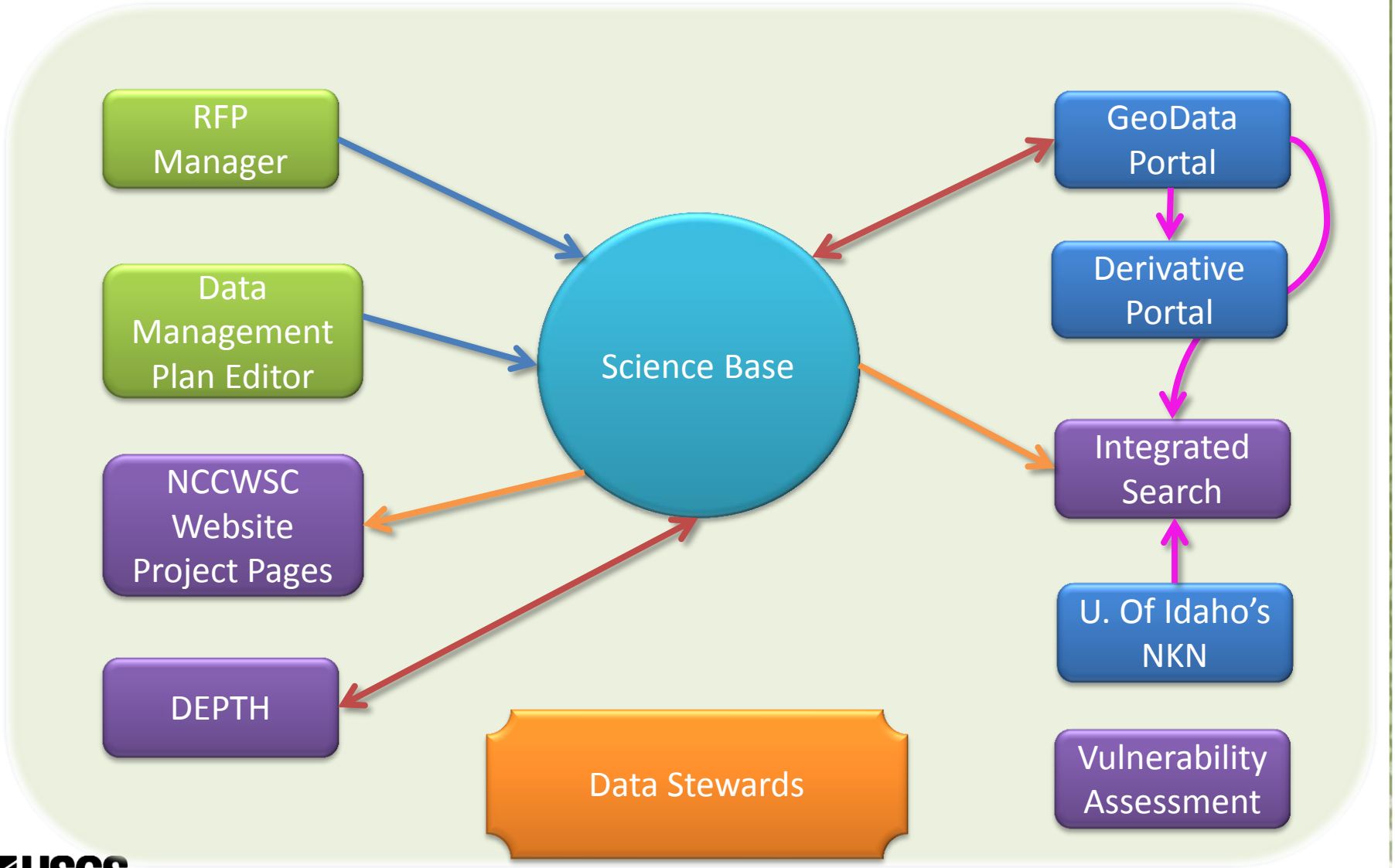
Data Policies

- All project products (data, models, etc.) will be shared (unless there is a good reason not to)
- Sharing happens when the project is complete
- Data management plans are required
- Common standards should be used
- Metadata must be provided

https://nccwsc.usgs.gov/content/data-policies-and-guidance

USGS science for a changing world

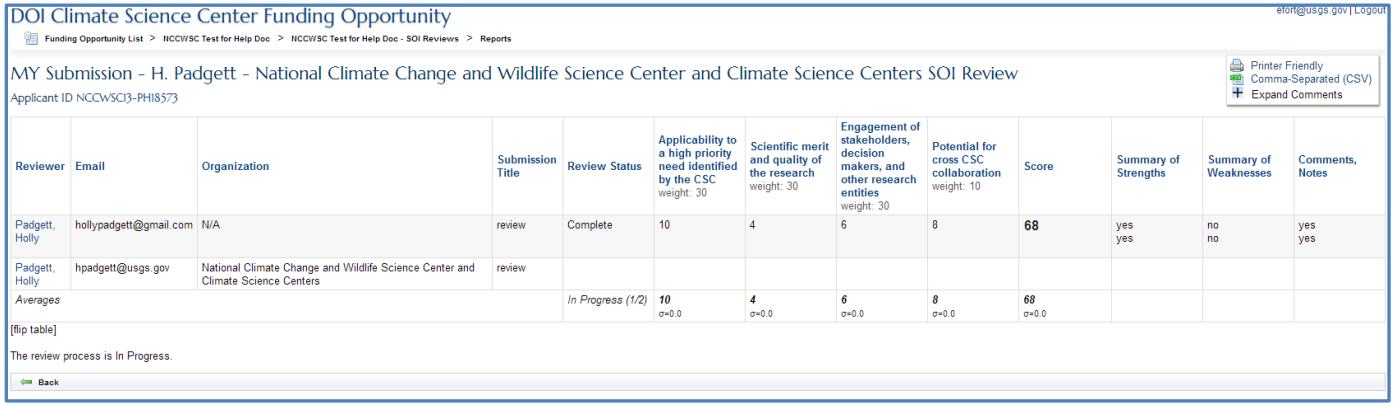
The Big Picture



USGS science for a changing world

RFP Manager

- Needed a way to collect proposal information and conduct peer review
- Ensures consistency in process and compliance with policies (including DMP)



DOI Climate Science Center Funding Opportunity

efort@usgs.gov | Logout

Funding Opportunity List > NCCWSC Test for Help Doc > NCCWSC Test for Help Doc - SOI Reviews > Reports

MY Submission - H. Padgett - National Climate Change and Wildlife Science Center and Climate Science Centers SOI Review

Applicant ID NCCWSC13-PH18573

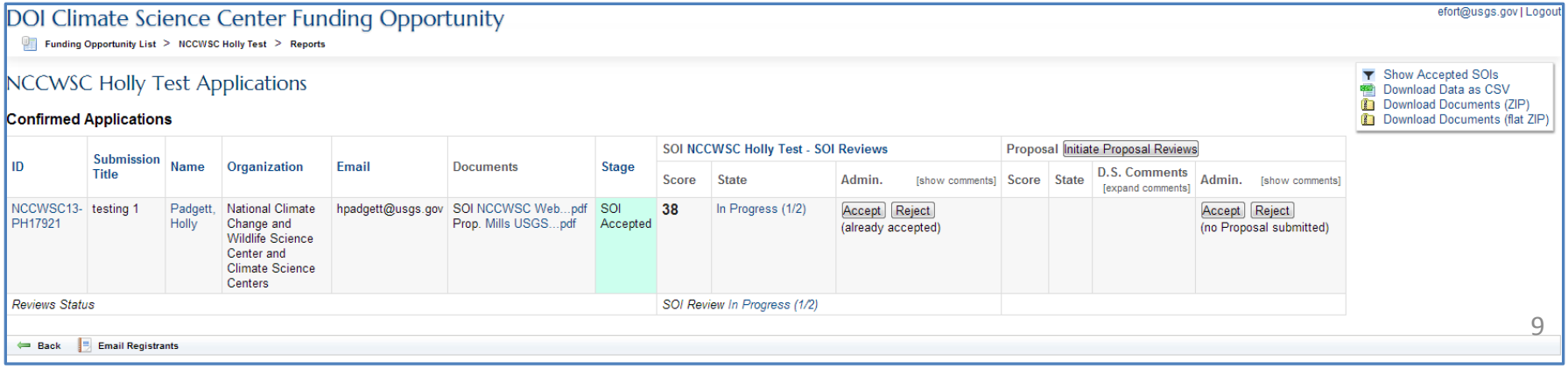
Printer Friendly
Comma-Separated (CSV)
Expand Comments

Reviewer	Email	Organization	Submission Title	Review Status	Applicability to a high priority need identified by the CSC weight: 30	Scientific merit and quality of the research weight: 30	Engagement of stakeholders, decision makers, and other research entities weight: 30	Potential for cross CSC collaboration weight: 10	Score	Summary of Strengths	Summary of Weaknesses	Comments, Notes
Padgett, Holly	hollypadgett@gmail.com	N/A	review	Complete	10	4	6	8	68	yes yes	no no	yes yes
Padgett, Holly	hpadgett@usgs.gov	National Climate Change and Wildlife Science Center and Climate Science Centers	review									
Averages				*In Progress (1/2)*	***10*** σ=0.0	***4*** σ=0.0	***6*** σ=0.0	***8*** σ=0.0	***68*** σ=0.0			

[flip table]

The review process is In Progress.

Back



DOI Climate Science Center Funding Opportunity

efort@usgs.gov | Logout

Funding Opportunity List > NCCWSC Holly Test > Reports

NCCWSC Holly Test Applications

Show Accepted SOIs
Download Data as CSV
Download Documents (ZIP)
Download Documents (flat ZIP)

Confirmed Applications

ID	Submission Title	Name	Organization	Email	Documents	Stage	SOI NCCWSC Holly Test - SOI Reviews			Proposal Initiate Proposal Reviews			
							Score	State	Admin. [show comments]	Score	State	D.S. Comments [expand comments]	Admin. [show comments]
NCCWSC13-PH17921	testing 1	Padgett, Holly	National Climate Change and Wildlife Science Center and Climate Science Centers	hpadgett@usgs.gov	SOI NCCWSC Web...pdf Prop. Mills USGS...pdf	SOI Accepted	**38**	In Progress (1/2)	Accept Reject (already accepted)				Accept Reject (no Proposal submitted)
Reviews Status							*SOI Review In Progress (1/2)*						

Back Email Registrants

Data Management Plan Template

- Organized by inputs and outputs
- For inputs, ask about new collections and use of existing data

1	[Name of Collection]
Description:	Describe the information that v landscape, etc.) of the data. Inc
Format:	Identify the formats in which th
Quality Checks:	Specify the procedures used to and an assessment of usability.
Source:	Identify the source for the data
Data Processing & Scientific Workflows:	Describe any data processing st manipulate the data, as approp
Backup & Storage:	Describe the approach for back project during the project.
Volume Estimate:	Estimate the volume of inform
Access & Sharing:	Prior to the completion of the p information/products and wha
Restrictions:	Identify any limitations on acce license restrictions, etc.) and pr documentation describing limit
Fees:	Identify any fees associated wit
Citation:	Provide citation for data produ

1	[Name of Output]
Description:	Describe the data output.
Data Management Resources:	Describe the proposal resources allocated for data management activities for the new data collected as a level of effort, total dollars allocated, or as a percentage of the total project's cost. Resources could include people's time or proposal funding.
Format:	Identify the formats in which the data will be generated, maintained, and made available.
Data Processing & Scientific Workflows:	Describe data processing steps or provide a scientific workflow you plan to use to manipulate the data, as appropriate.
Quality Checks:	Specify the procedures for ensuring data quality during the project.
Metadata:	Identify the metadata standard that will be used to describe the data and products (FGDC, ISO, EML, etc.)
Volume Estimate:	Estimate the volume of information generated: megabyte (MB), GB, TB, or PB.
Backup & Storage:	Describe the approach for backup and storage of the information associated with the research project during the project.
Repository for Data:	In addition to the NCCWSC repository (ScienceBase), identify any other repositories where you plan to share your data.
Access & Sharing:	Prior to the completion of the project, specify who should have access to project information/products and what type of access (Public, Read, Write, No Access).
Exclusive Use:	Project data and associated products should be available publically at the end of the project. If a request to limit access for a period of time after project completion is needed, please identify the length of time and the reason for the extension. (Request cannot be more than two years.)
Restrictions:	Identify any limitations on access or reuse (e.g., sensitive data, restricted data, software with license restrictions, etc.) and provide justification for restriction. Provide citation or documentation describing limitations if due to policies or legal reasons.
Citation:	Specify how the project's data should be cited.
Digital Object Identifier (DOI)/Link:	Provide a digital object identifier (DOI)/link to the project when available publically.
Contact:	Provide a point(s) of contact if questions arise related to the data and associated products (name, email, and phone number).

Data Management Plan Review

- For each proposal, the proposal data management plan (DMP) is reviewed and comments are provided
- For funded proposals, a data steward works with the research team to complete the full DMP
- At project completion, the data steward works with the research team to transfer the products to the NCCWSC repository - ScienceBase

USGS
science for a changing world

ScienceBase

What is it?

- Website
- Provides a data cataloging and collaborative data management platform for USGS scientists and partners
- Central search and discovery application along with web services that facilitate other applications
- Open source project
- Bigger than just the CSCs

https://www.sciencebase.gov/catalog/

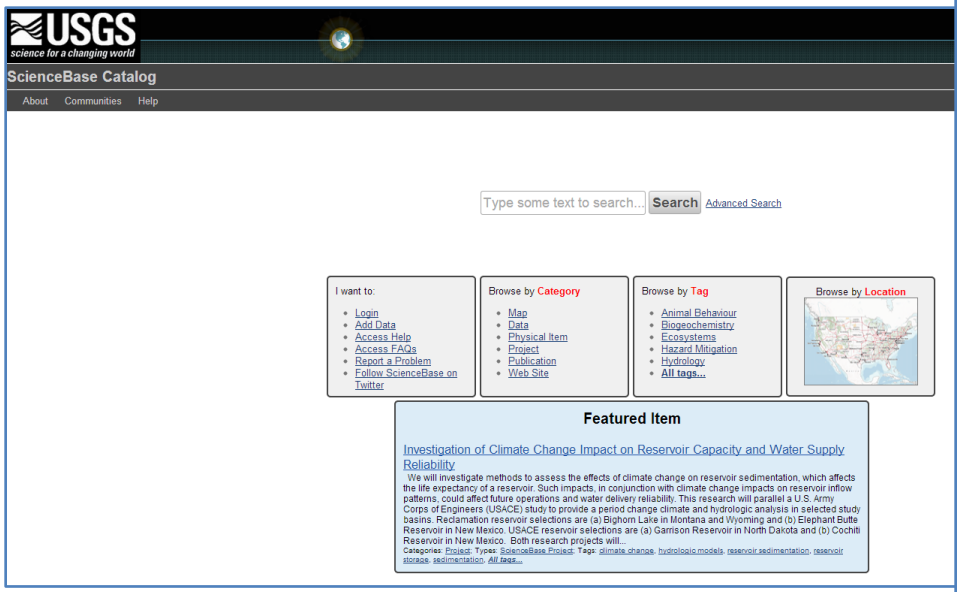
Why ScienceBase?

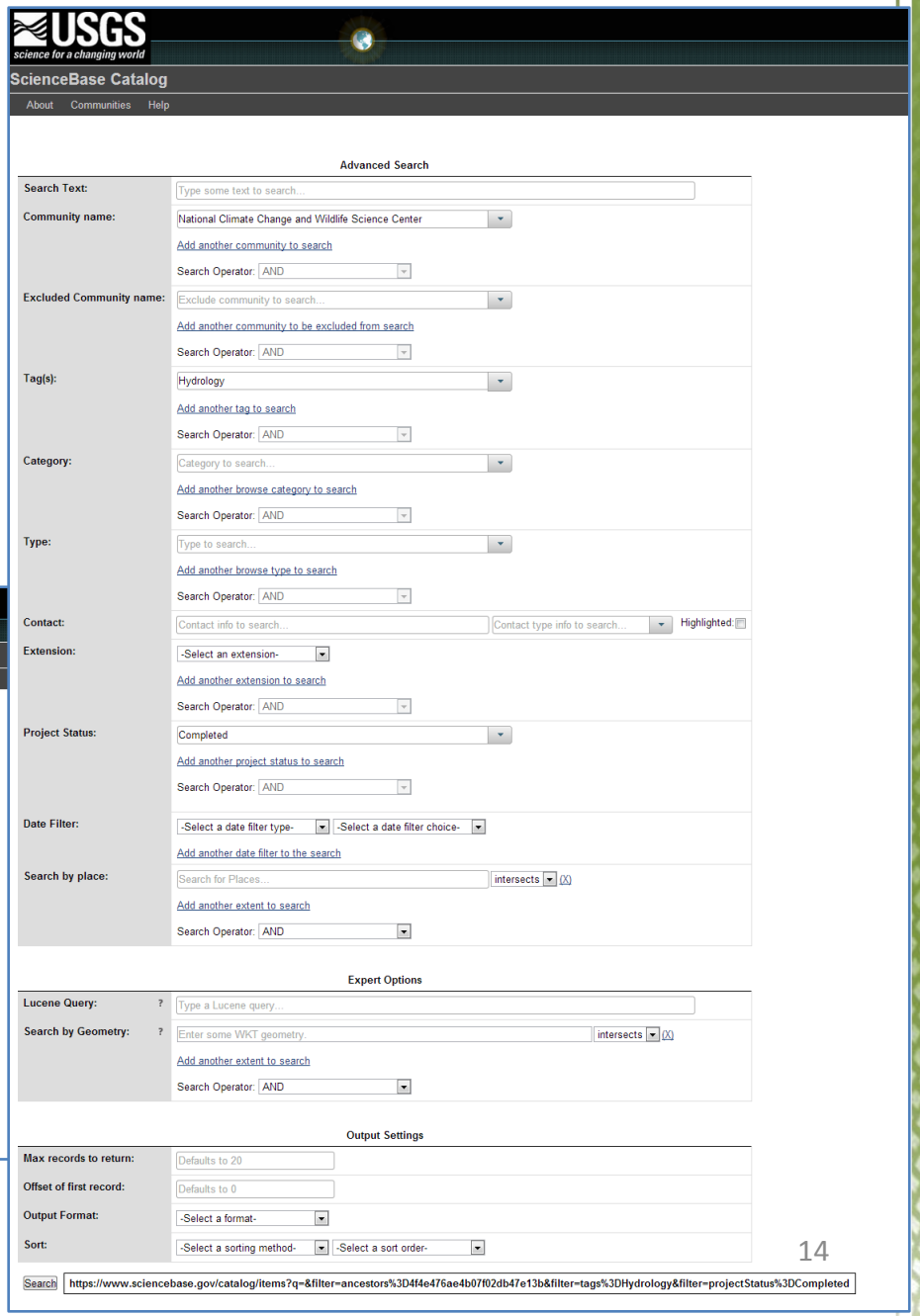
- Powerful and flexible
- Can store many different kinds of data – including large datasets and models
- Core fields and extensible
- Searchable
- Geospatial component
- Available to non-DOI users
- Information can be open (public) or restricted
- Meets federal IT and security requirements
- Web services

USGS
science for a changing world

Finding "Stuff" in ScienceBase

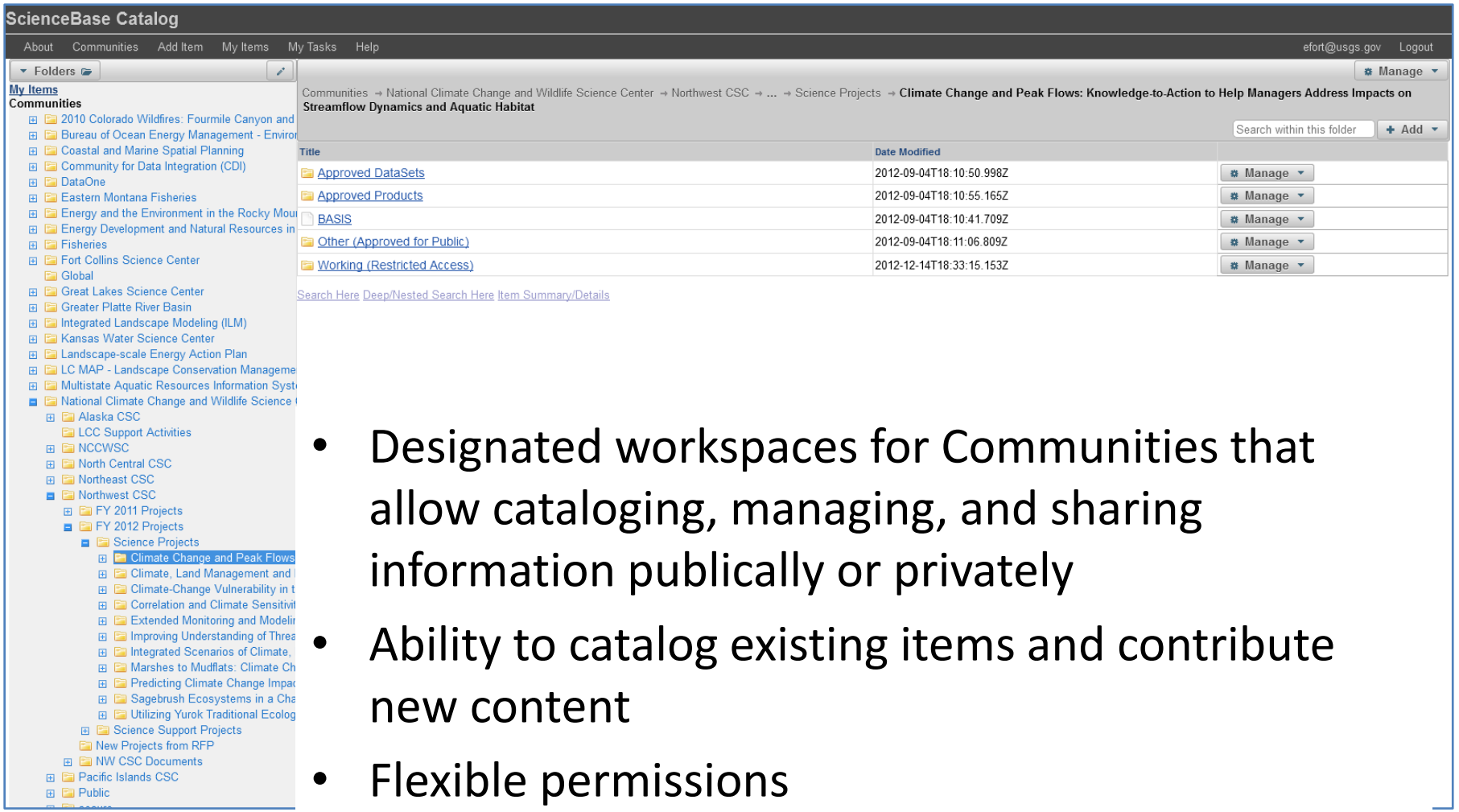
- Lots of ways to search





USGS
science for a changing world

ScienceBase Organization



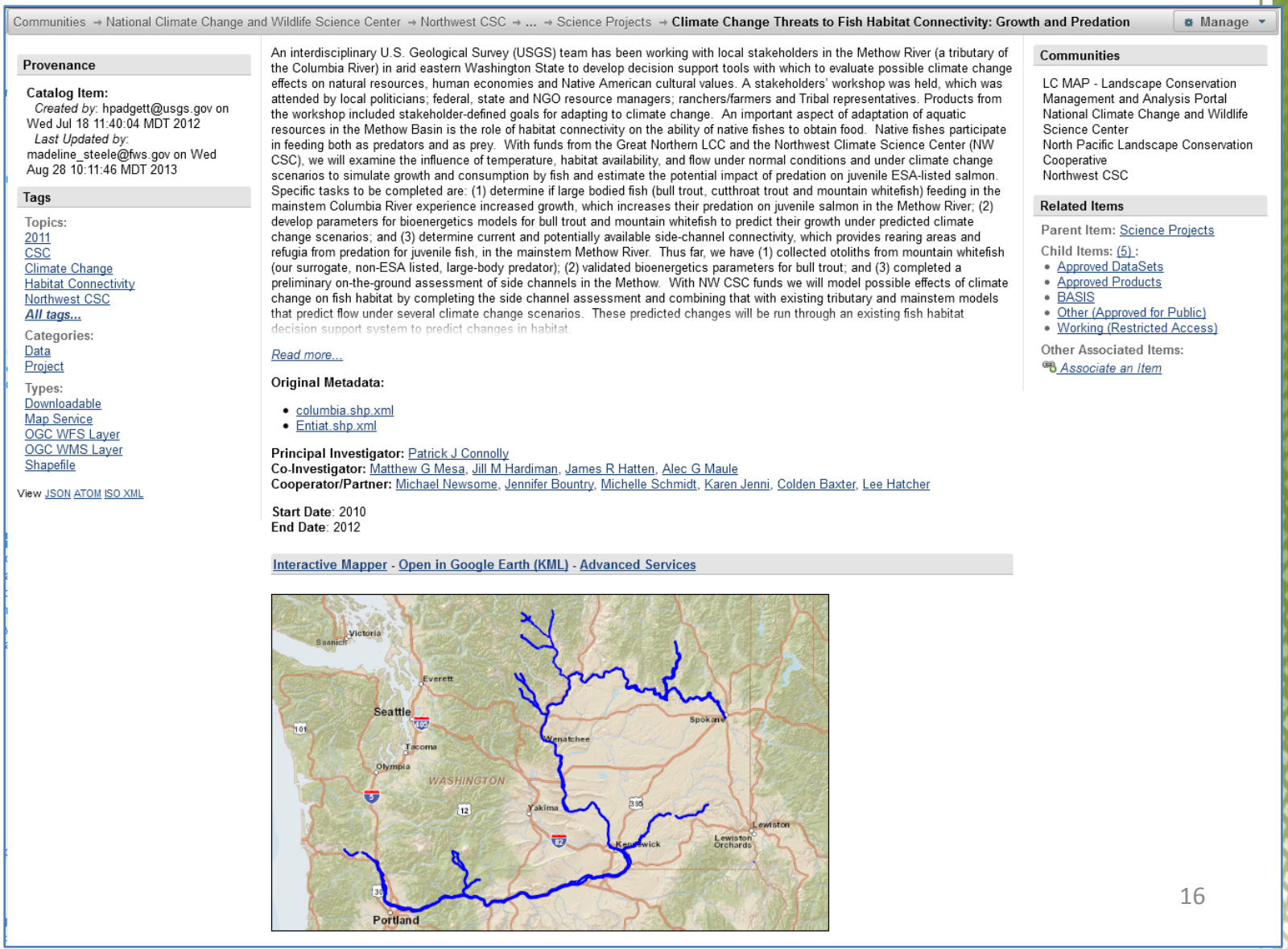
- Designated workspaces for Communities that allow cataloging, managing, and sharing information publically or privately
- Ability to catalog existing items and contribute new content
- Flexible permissions
- Collections organized by datasets, products, other, and a working folder

USGS science for a changing world

ScienceBase Project Record

Geospatial
- Shapefiles
- Geotiffs
- Polygon

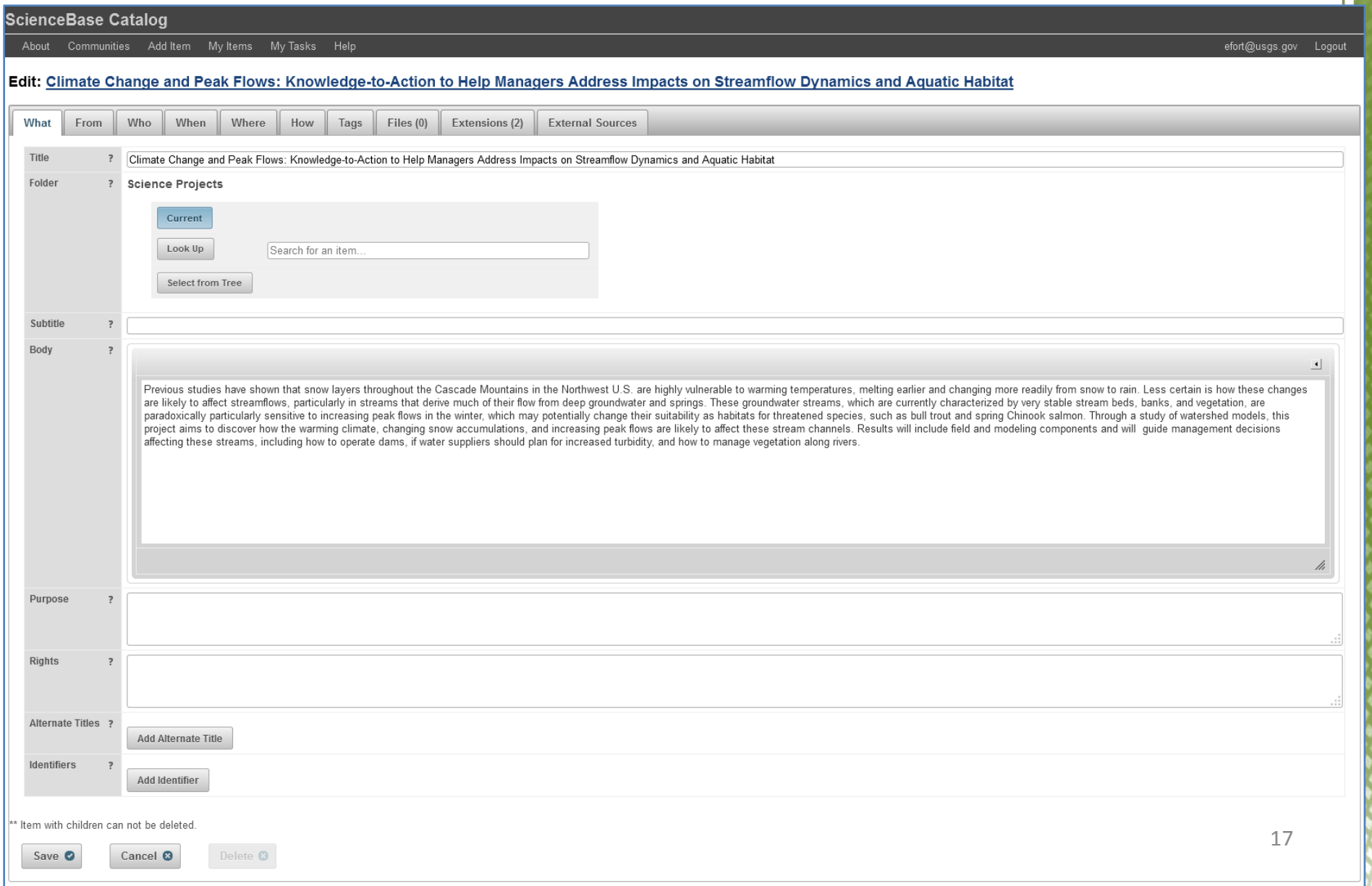
Note
- Tags
- Commun-itities
- Related items



USGS science for a changing world

Editing a Project's Metadata

- Extensions added for custom fields

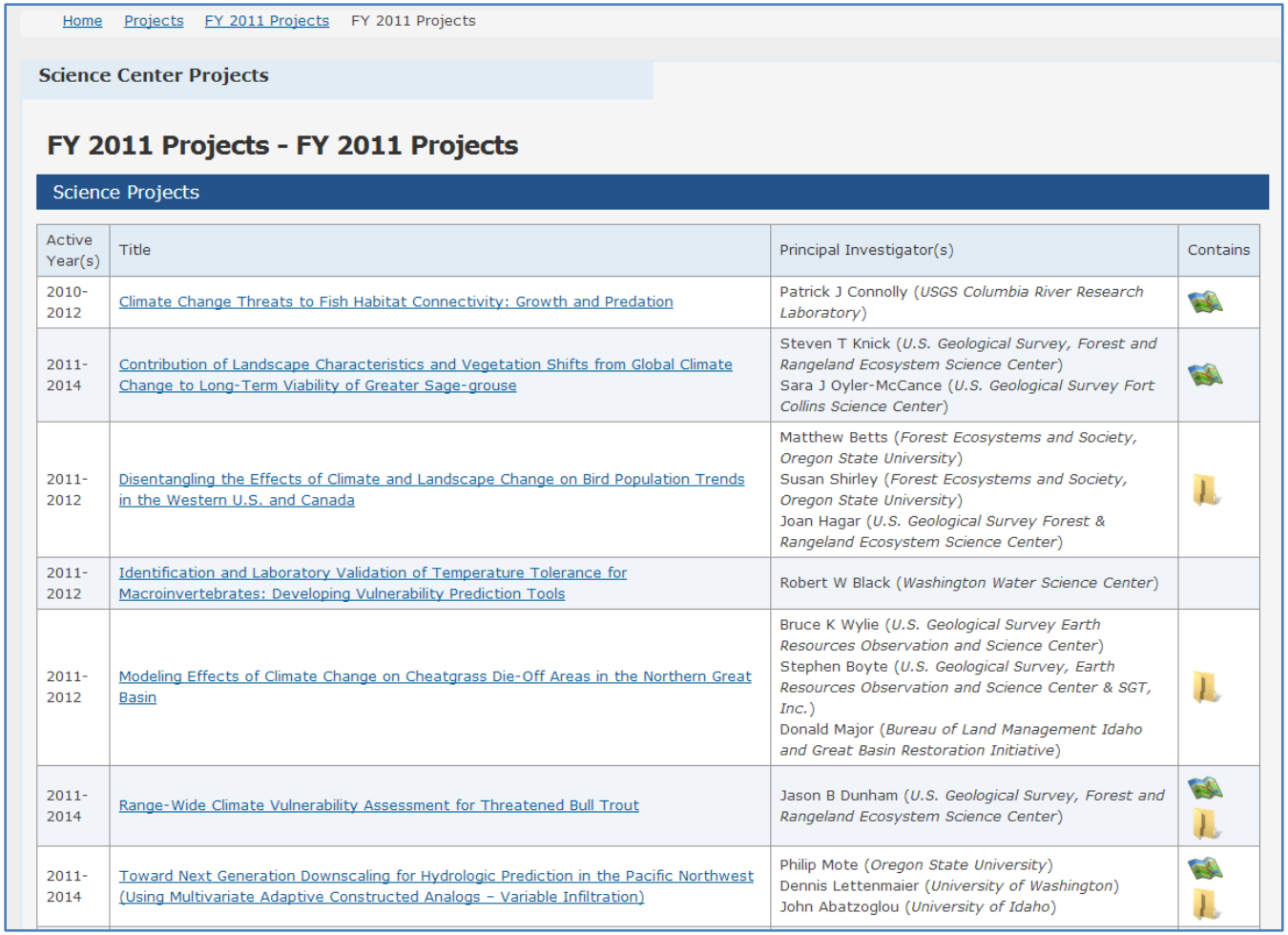


ScienceBase Web Services

- Other sites can consume and present information from ScienceBase with their own look and feel using ScienceBase's REST service architecture using JSON
- Examples:
  - The NCCWSC website project pages consume ScienceBase web services and styled by Drupal
  - DEPTH consumes ScienceBase web services and filters for records with project tag

NCCWSC Website Project Pages - Summary

- Organized by CSC and by Fiscal Year Funded
- Icon to indicate map or data/product



Home Projects FY 2011 Projects FY 2011 Projects

Science Center Projects

FY 2011 Projects - FY 2011 Projects

Science Projects

Active Year(s)	Title	Principal Investigator(s)	Contains
2010-2012	Climate Change Threats to Fish Habitat Connectivity: Growth and Predation	Patrick J Connolly (*USGS Columbia River Research Laboratory*)	
2011-2014	Contribution of Landscape Characteristics and Vegetation Shifts from Global Climate Change to Long-Term Viability of Greater Sage-grouse	Steven T Knick (*U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center*) Sara J Oyler-McCance (*U.S. Geological Survey Fort Collins Science Center*)	
2011-2012	Disentangling the Effects of Climate and Landscape Change on Bird Population Trends in the Western U.S. and Canada	Matthew Betts (*Forest Ecosystems and Society, Oregon State University*) Susan Shirley (*Forest Ecosystems and Society, Oregon State University*) Joan Hagar (*U.S. Geological Survey Forest & Rangeland Ecosystem Science Center*)	
2011-2012	Identification and Laboratory Validation of Temperature Tolerance for Macroinvertebrates: Developing Vulnerability Prediction Tools	Robert W Black (*Washington Water Science Center*)	
2011-2012	Modeling Effects of Climate Change on Cheatgrass Die-Off Areas in the Northern Great Basin	Bruce K Wylie (*U.S. Geological Survey Earth Resources Observation and Science Center*) Stephen Boyte (*U.S. Geological Survey, Earth Resources Observation and Science Center & SGT, Inc.*) Donald Major (*Bureau of Land Management Idaho and Great Basin Restoration Initiative*)	
2011-2014	Range-Wide Climate Vulnerability Assessment for Threatened Bull Trout	Jason B Dunham (*U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center*)	
2011-2014	Toward Next Generation Downscaling for Hydrologic Prediction in the Pacific Northwest (Using Multivariate Adaptive Constructed Analogs – Variable Infiltration)	Philip Mote (*Oregon State University*) Dennis Lettenmaier (*University of Washington*) John Abatzoglou (*University of Idaho*)	

CSC Projects: https://nccwsc.usgs.gov/project-pages/4f4e476ae4b07f02db47e13b

USGS science for a changing world

NCCWSC Website Project Pages - Detail

Powered by ScienceBase

Modeling Effects of Climate Change on Cheatgrass Die-Off Areas in the Northern Great Basin

Project Information

Principal Investigator(s):
Bruce K Wylie (*U.S. Geological Survey Earth Resources Observation and Science Center*)
Stephen Boyte (*U.S. Geological Survey, Earth Resources Observation and Science Center & SGT, Inc.*)
Donald Major (*Bureau of Land Management Idaho and Great Basin Restoration Initiative*)

Start Date: October 2011

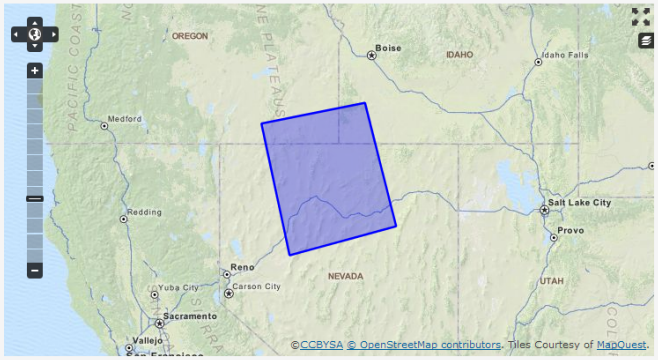
End Date: September 2012

Project Status: Completed

Tags: Climate Change, Cheatgrass, Great Basin, CSC, Northwest CSC, 2011, Science Project

CSC/NCCWSC Affliation: Northwest CSC

Fiscal Year: FY 2011 Projects



Summary

Cheatgrass (*Bromus tectorum*) is a dominant invasive species across large areas of the Great Basin. In recent years, the die-off of cheatgrass has been observed across relatively large areas in the region with an estimated 500,000 acres of affected area reported in the general vicinity of Winnemucca, NV. However, actual extent of the phenomenon could be considerably larger as die-offs are occurring in smaller areas across portions of the Northern Great Basin. As part of the Bureau of Land Management's (BLM) Integrated Cheatgrass Dieoff Project, U.S. Geological Survey (USGS) Earth Resources Observation Systems (EROS) Center scientists in collaboration with Don Major, BLM Landscape Ecologist, have developed a cheatgrass performance model that incorporates seasonally integrated normalized difference vegetation index (NDVI) from the enhanced Moderate Resolution Imaging Spectroradiometer (eMODIS) along with environmental attributes. Based on the die-off areas in the area surrounding Winnemucca and in the Owyhee Uplands, we propose to predict areas of potential cheatgrass die-offs under future climate projections and make climate-based forecasts of these die-off areas.

Products & Data

Cheatgrass Dieoff Time-series Dynamics (2000-2010)
Land Cover Applications and Global Change (External URL)

Cheatgrass dieoff in Northern Great Basin Final Report
Cheatgrass dieoff in Northern Great Basin FINAL REPORT 09JUL12.pdf (Download)

Identifying cheatgrass dieoff in the Great Basin by integrating eMODIS NDVI data with ecological models
Society for Range Management (External URL)

Mapping interannual cheatgrass production and dieoff in the Great Basin using remote sensing data and ecological models
YouTube Video (External URL)

ReadMe file to accompany the report "Mapping Cheatgrass Dieoff in the Northern Great Basin using Ecosystem Performance Modeling"
ReadMe.docx (Download)

Data
- Cheatgrass Percent Cover Maps
- Cheatgrass Dieoffs

Maps
- Cheatgrass PercentCover
- Cheatgrass Dieoffs

ScienceBase Url: https://www.sciencebase.gov/catalog/item/5006f498e4b0abf7ce733f92

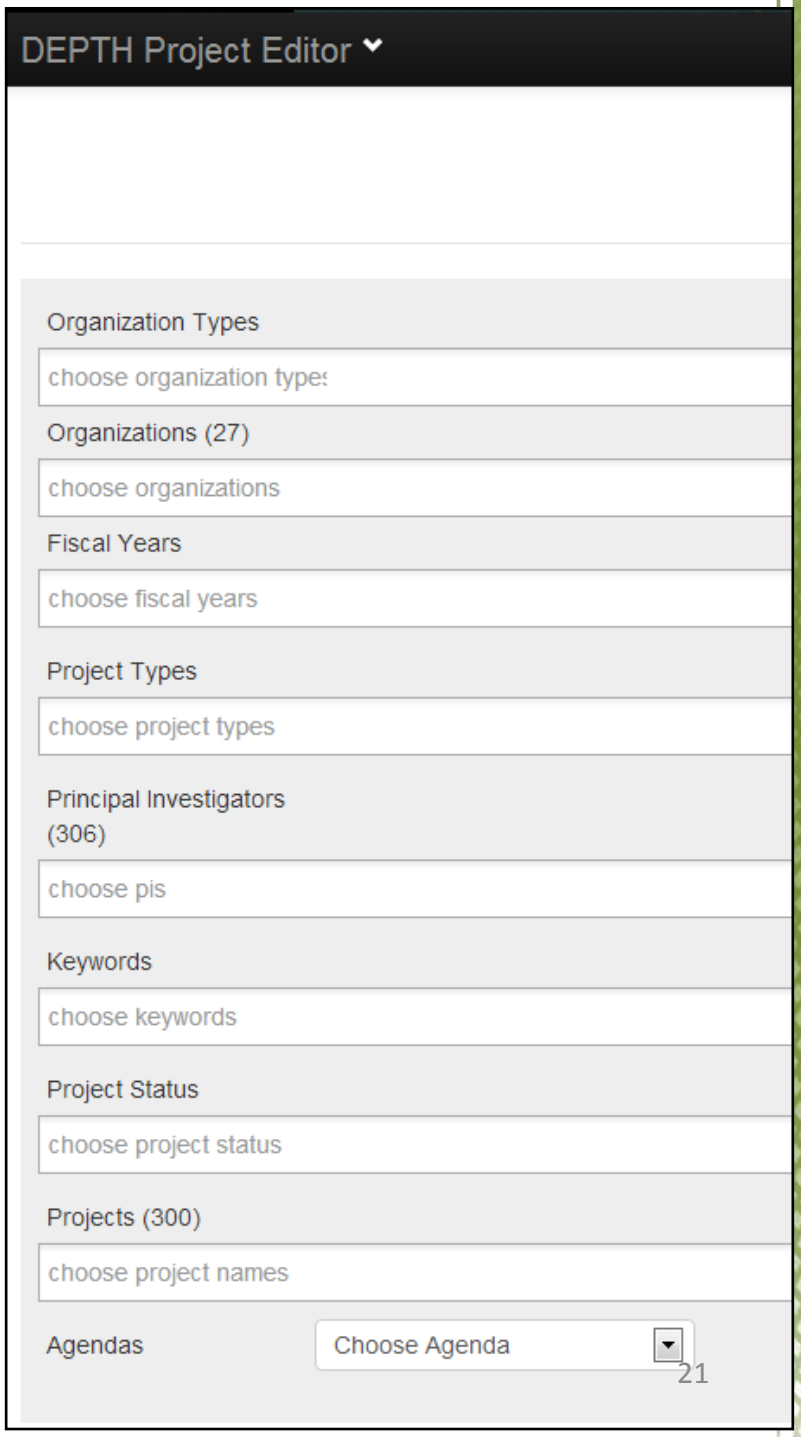
Example URL: https://nccwsc.usgs.gov/display-project/5006c1e5e4b0abf7ce733f3b/5006f498e4b0abf7ce733f92

USGS science for a changing world

DEPTH

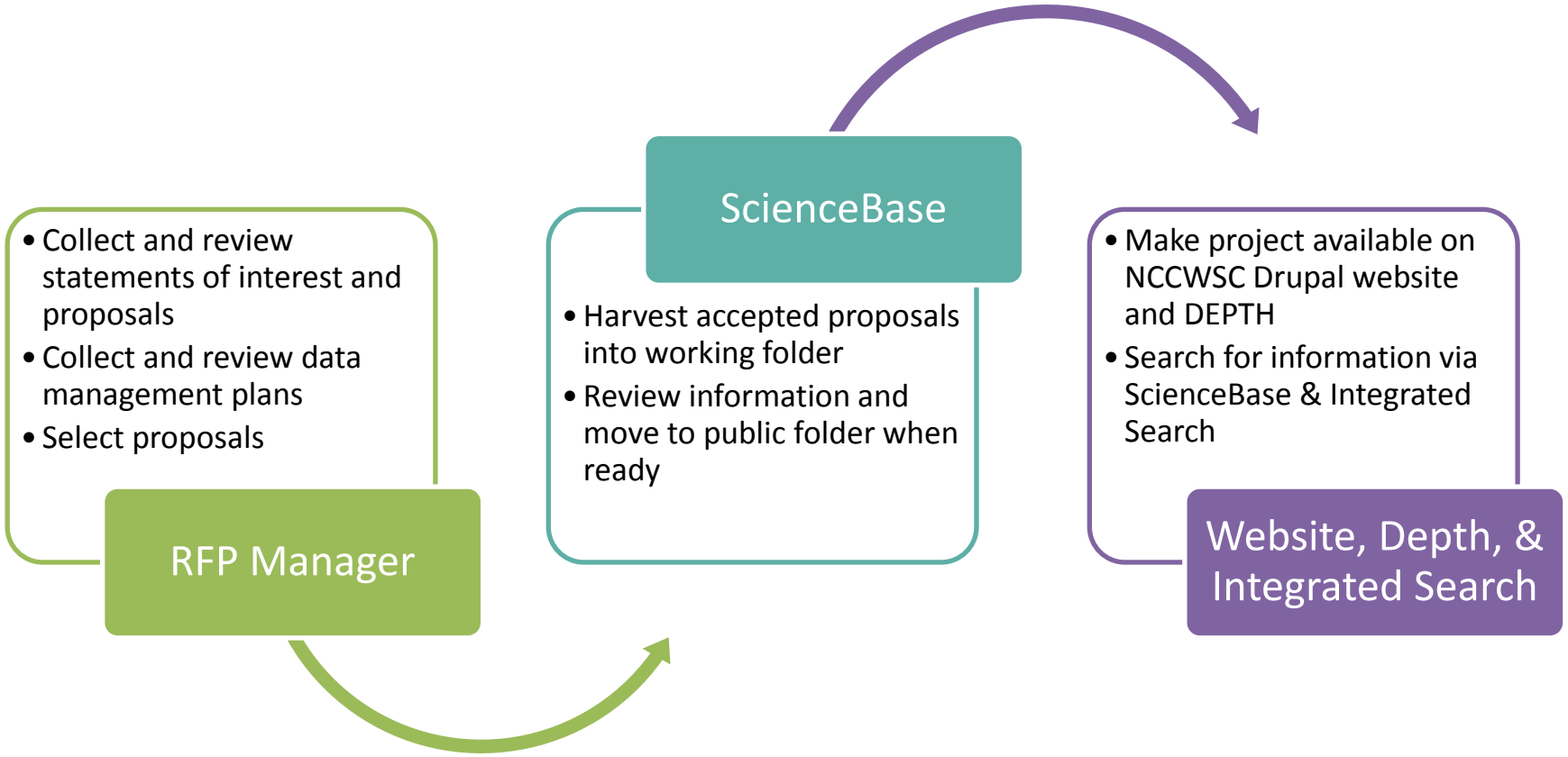
- Project-centric view of ScienceBase information
- Searchable by many filters, including science agenda
- Intuitive entry of new records
- Assists with CSC regional coordination

Example:
https://my.usgs.gov/depth/#/viewProject/5006f498e4b0abf7ce733f92/csc



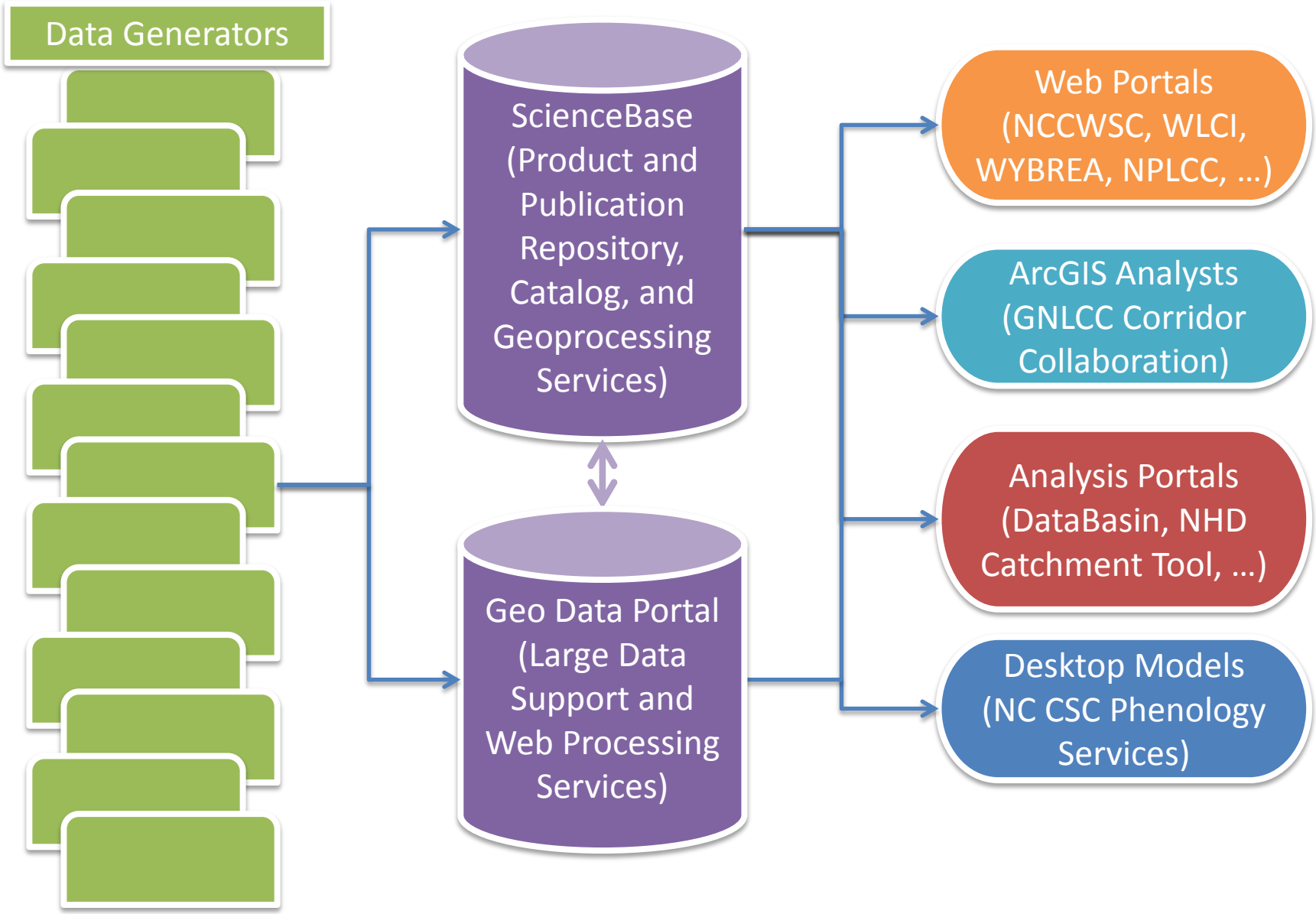
USGS science for a changing world

Steps Along the Way



USGS science for a changing world

How Does This Work with Other Stuff?



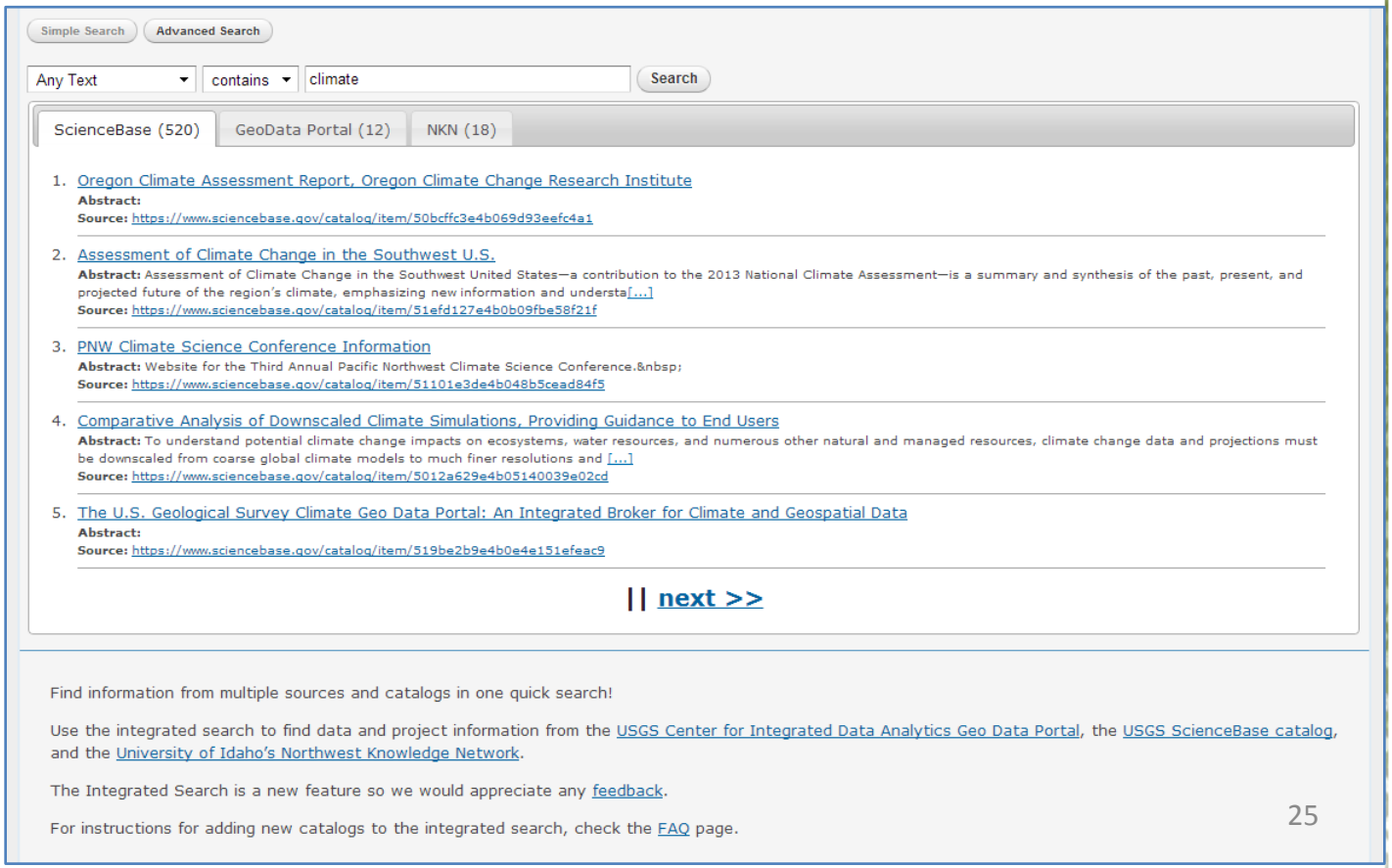
Geo Data Portal

- Tool to help scientists use and subset large datasets such as climate information
- Standards based
- Integrated with ScienceBase

For more information, see the Geo Data Portal webinar at https://nccwsc.usgs.gov/webinar/114

Integrated Search

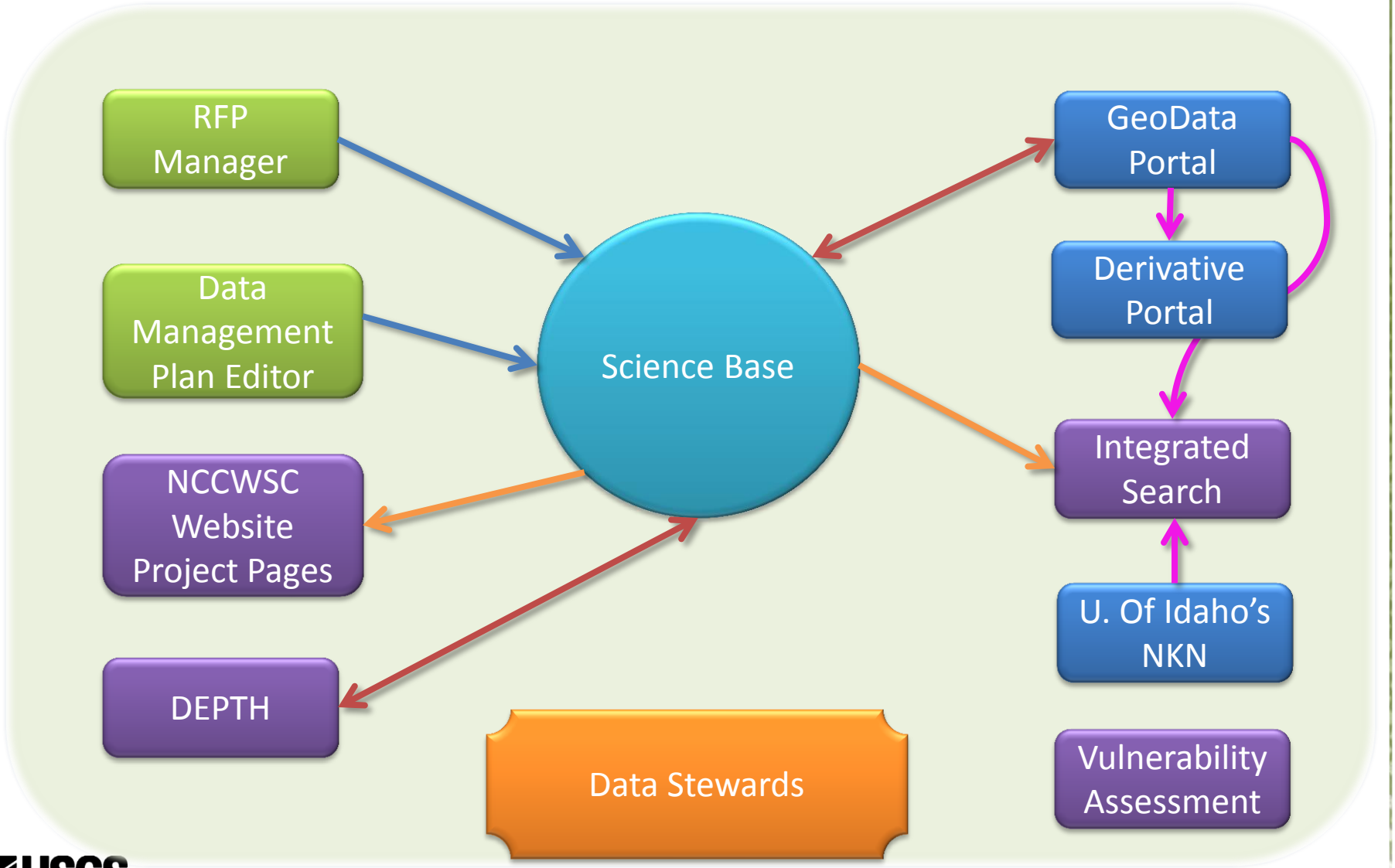
- Search multiple data repositories at the same time
- Uses common metadata standard
- Works by using csw (catalog service for the web)search
- Give us feedback!



https://nccwsc.usgs.gov/integrated_search

USGS science for a changing world

The Big Picture



USGS science for a changing world

Link to LCCs and Others

- ScienceBase and related tools used or being considered by several LCCs (aka LC Map)
- ScienceBase team part of the LCC Network's Integrated Data Management Network project (IDMN)
- ScienceBase also used by other parts of USGS and other partners
- Cross-fertilization of data management working groups

USGS science for a changing world

LANDSCAPE CONSERVATION COOPERATIVES

So What Does All This Accomplish?

- Provides core, basic capabilities for CSCs that also provide necessary information to the NCCWSC
  - RFP Manager
  - Project Information
  - Data Management
  - Data Repository
- Leverages ongoing work and developed features and tools that others can use
- Supports growth and program's future
- Supports good data management

USGS
science for a changing world

Where Do We Go From Here?

- Learn lessons from DMPs to improve guidance, clarify, walk line of getting the right amount of information at the proposal stage
- Identify areas for future collaboration and integration with partners and CSC members
- Develop additional tools and features as needed (and as resources allow)
- Add features (DOIs, better metadata tools)
- Follow evolution of government policy and tools

USGS
science for a changing world

Questions?

Emily Fort

efort@usgs.gov

703-648-4082

USGS science for a changing world